BEST AVAILABLE COPY

IB04/03618

Europäisches Patentamt | European Patent Office | Office européen des brevets

REC'D 11 NOV 2004
WIPO PCT

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. | Patent application No. | Demande de brevet n°

03300205.6

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 03300205.6
Demande no:

Anmeldetag:
Date of filing: 07.11.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Encoding method and device

In Anspruch genommene Prioriät(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT RO SE SI SK TR LI

ENCODING METHOD AND DEVICE

FIELD OF THE INVENTION

The present invention relates to a video encoding method provided for encoding an input image sequence consisting of successive groups of frames in which each frame is itself subdivided into blocks, said method comprising the steps of :

- preprocessing said sequence on the basis of a so-called content-change strength (CCS) computed for each frame by applying some predetermined rules ;
- estimating a motion vector for each block of the frames ;
- generating a predicted frame using said motion vectors respectively associated to the blocks of the current frame ;
- applying to a difference signal between the current frame and the last predicted frame a transformation sub-step producing a plurality of coefficients and followed by a quantization sub-step of said coefficients ;
- coding said quantized coefficients.

Said invention, applicable to video encoding devices that require reference frames for reducing e.g. temporal redundancy, like motion estimation and compensation devices, may be used for instance in devices like digital video cameras, mobile phones or digital video recording devices. Furthermore, applications for coding or transcoding video can be enhanced using the technique according to the invention.

BACKGROUND OF THE INVENTION

In video compression, low bit rates for the transmission of a coded video sequence may be obtained by (among others) a reduction of the temporal redundancy between successive pictures. Such a reduction is based on motion estimation (ME) and motion compensation (MC) techniques. Performing ME and MC for the current frame of the video sequence however requires reference frames. Taking MPEG-2 as an example, different frames types, namely I-, P- and B-frames, have been defined, for which ME and MC is performed differently : I-frames (or intra frames) are independently coded, without any reference to past or future frames (no ME and MC is performed), while P-frames (or forward predicted pictures) are encoded relatively to past frames and B-frames (or bidirectional predicted frames) are encoded relatively to two reference frames (a past frame and a future frame). The I- and P-frames serve as reference frames.

In order to obtain good frame predictions, these reference frames need to be of high quality, i.e. many bits have to be spent to code them, whereas non-reference frames can be of lower quality (for this reason, a higher number of non-reference frames, B-frames in the case of MPEG-2, generally lead to lower bit rates). In order to indicate which input frame is processed as an I-frame, a P-frame or a B-frame, a structure based on groups of pictures (GOPs) is defined in MPEG-2. More precisely, a GOP uses two parameters N and M, where N is the temporal distance between two I-frames and M is the temporal distance between reference frames. For example, an (N,M)-GOP with N=12 and M=4 is commonly used, defining an "I B B B P B B B P B B B" structure.

Succeeding frames generally have a higher temporal correlation than frames having a larger temporal distance between them. Therefore shorter temporal distances between the reference and the currently predicted frame on the one hand lead to higher prediction quality, but on the other hand imply that less non-reference frames can be used. Both the higher prediction quality and a higher number of non-reference frames generally result in lower bit rates, but they work against each other since the frame prediction quality results from shorter temporal distances only.

However, said quality also depends on the usefulness of the reference frames to actually serve as references. For example, it is obvious that with a reference frame located just before a scene change, the prediction of a frame located just after the scene change is not possible with respect to said reference frame, although they may have a frame distance of only 1. One the other hand, in scenes with a steady or almost steady content (like video conferencing or news), even a frame distance of more than 100 can still result in high quality prediction.

From the above-mentioned examples, it appears that a fixed GOP structure as the commonly used (12, 4)-GOP may be inefficient for coding a video sequence, because reference frames are introduced too frequently, in case of a steady content, or at a unsuitable position, if they are located just before a scene change. Scene-change detection is a known technique that can be exploited to introduce an I-frame at a position where a good prediction of the frame is not possible due to a scene change, but some sequences do not profit from such techniques if the frame content is almost completely different after some frames having high motion, with however no scene change at all (for instance, in a sequence where a tennis player is continuously followed within a single scene). A previous European patent application, already filed by the applicant on October 14, 2003, with the filing number

03300155.3 (PHFR030124) has then described a new method for finding better reference frames.

SUMMARY OF THE INVENTION

It is the object of the invention to propose a method based on said previous method for finding good reference frames, but allowing to reduce more noticeably the coding cost.

To this end, the invention relates to a video encoding method such as defined in the introductory paragraph of the description and in which said CCS is used in said quantization sub-step for modifying the quantization factor, said CCS and the quantization factor increasing or decreasing simultaneously.

The invention also relates to a device for implementing said method.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described, by way of example, with reference to the accompanying drawings in which :

- Fig. 1 illustrates the rules used, according to the description given in the previous European patent application cited above, for defining the place of the reference frames of the video sequence to be coded ;
- Fig.2 shows an encoder carrying out the method described in said previous European patent application in the MPEG encoding case ;
- Fig.3 shows an encoder carrying out the method according to the invention.

DETAILED DESCRIPTION OF THE INVENTION

The document mentioned above describes a method for finding which frames in the input sequence can serve as reference frames, in order to reduce the coding cost. The principle of this method is to measure the strength of content change on the basis of some simple rules, such as listed below and illustrated in Fig.1 : the measured strength of content change is quantized to levels (for instance 5, said number being however not a limitation), and I-frames are inserted at the beginning of a sequence of frames having content-change strength (CCS) of level 0, while P-frames are inserted before a level increase of CCS occurs or after a level decrease of CCS occurs. The measure may be for instance a simple block classification that detects horizontal and vertical edges, or other types of measures based on luminance, motion vectors, etc.

An implementation of this previous method in the MPEG encoding case is described in Fig.2. The encoder comprises a coding branch 101 and a prediction branch 102. The signals to be coded, received by the branch 101, are transformed into coefficients in a DCT module 11 and quantized in a quantization module 12, the quantized coefficients being then coded in a coding module 13, together with motion vectors MV. The prediction branch, receiving as input signals the signals available at the output of the quantization module 12, comprises in series an inverse quantization module 21, an inverse DCT module 22, an adder 23, a frame memory 24, an MC circuit 25 and a subtracter 26. The MC circuit 25 also receives the motion vectors MV generated by a ME circuit 27 (many types of motion estimators may be used) from the input reordered frames (defined as explained below) and the output of the frame memory 24, and these motion vectors are also sent towards the coding module 13, the output of which ("MPEG output") is stored or transmitted in the form of a multiplexed bitstream.

The video input of the encoder (successive frames Xn) is preprocessed in a preprocessing branch 103. First a GOP structure defining circuit 31 defines from the successive frames the structure of the GOPs. Frame memories 32a, 32b, ... are then provided for reordering the sequence of I, P, B frames available at the output of the circuit 31 (the reference frames must be coded and transmitted before the non-reference frames depending on said reference frames). These reordered frames are sent on the positive input of the subtracter 26 (the negative input of which receives, as described above, the output predicted frames available at the output of the MC circuit 25, these output predicted frames being also sent back to a second input of the adder 23). The output of the subtracter 26 delivers frame differences that are the signals processed by the coding branch 101. For the definition of the GOP structure, a CCS computation circuit 33 is provided.

It has then been observed that the higher the CCS – which can result from motion – the less the viewer can really follow the presented video. It is consequently proposed, according to the present invention, to increase or decrease the quantization factor used in the module 12 as a function of the CCS - said CCS and the quantization factor increasing or decreasing simultaneously – which can be obtained by sending the output information of the CCS computation circuit towards the quantization module 12 of the coding branch. It is known, indeed, that the coding module 13 comprises in fact, in series, a variable-length coding (VLC) circuit 131 and a buffer memory 132, the output of said memory being sent back towards a rate control circuit 133 for modifying the quantization factor.

According to the invention, and as shown in Fig.3 in which similar circuits are designated by the same references as in Fig.2, an additional connection 200 intended to allow to implement the proposed modification of quantization factor is provided between the CCS computation circuit 33 and the rate control circuit 133 and also between said circuit 33 and the quantization module 12 of the coding branch. This connection 200 extends two coding modes of the coding system, namely a so-called open-loop coding mode (without bit-rate control) and a closed loop coding mode (with bit-rate control).

In the open-loop coding mode for example, the quantizer settings are usually fixed. The resulting bit rate of the encoded stream is automatically lower for simple scenes (less residue needs to be coded) than for complex scenes (higher residue needs to be coded). Coding cases as described above, where the sequence contains high motion, result in complex scenes that are coded with high bit-rates. The bit-rates for the high-motion scenes can be reduced by higher quantization, thereby removing spatial details of these scenes that the observer cannot follow due to the motion. The quantization can be controlled by defining a quantization factor *q_ccs* which is a function of CCS and the original fixed quantizer factor, called *q_fixed* :

$$q\_ccs = q\_fixed + f(CCS),$$

where $f()$ is a function resulting in positive integers 0......($q\_max - q\_fixed$) to increase *q_ccs* from *q_fixed* upto an allowed maximum *q_max*. Examples for $f()$ are $f1(CCS) = round\ (\ CCS * (q\_max - q\_fixed)\ /\ (CCS\_max)\ )$ or $f2(CCS) = round\ (\ (q\_max - q\_fixed + 1)\ \hat{}\ (CCS/CCS\_max) - 1)$ for CCS=0..*CCS_max*.

In closed-loop coding, the quantization factor *q_adapt* is adapted in order to achieve a desired predefined bit rate. Bit-rate controllers that are required for closed-loop coding work basically with bit budgets and chose *q_adapt* based on the available budget. This means that the quantization factor *q_css* as described for open-loop coding can be used, and only *q_fixed* has to be replaced with *q_adapt*. Then, compared to an unmodified rate controller, the bit budget will increase with higher CCS, and these additional bits are automatically spent on frames with lower CCS, because the *q_adapt* value will decrease due to the increased bit budget.

CLAIMS :

1. A video encoding method provided for encoding an input image sequence consisting of successive groups of frames in which each frame is itself subdivided into blocks, said method comprising the steps of :

- preprocessing said sequence on the basis of a so-called content-change strength (CCS) computed for each frame by applying some predetermined rules ;
- estimating a motion vector for each block of the frames ;
- generating a predicted frame using said motion vectors respectively associated to the blocks of the current frame ;
- applying to a difference signal between the current frame and the last predicted frame a transformation sub-step producing a plurality of coefficients and followed by a quantization sub-step of said coefficients ;
- coding said quantized coefficients ;

wherein said CCS is used in said quantization sub-step for modifying the quantization factor, said CCS and the quantization factor increasing or decreasing simultaneously.

2. A video encoding device provided for encoding an input image sequence consisting of successive groups of frames in which each frame is itself subdivided into blocks, said device comprising the following means :

- preprocessing means, provided for preprocessing said sequence on the basis of a so-called content-change strength (CCS) computed for each frame by applying some predetermined rules ;
- estimating means, provided for estimating a motion vector for each block of the frames ;
- generating means, provided for generating a predicted frame on the basis of said motion vectors respectively associated to the blocks of the current frame ;
- transforming and quantizing means, provided for applying to a difference signal between the current frame and the last predicted frame a transformation producing a plurality of coefficients and followed by a quantization of said coefficients ;
- coding means, provided for encoding said quantized coefficients ;

wherein an output of said preprocessing means is received on an input of said transformation and quantization means for modifying the quantization factor on the basis of said CCS, said CCS and the quantization factor increasing or decreasing simultaneously.

Abstract

The invention relates to a video encoding method provided for encoding an input image sequence consisting of successive groups of frames in which each frame is itself subdivided into blocks, and to the corresponding video encoding device. This method and device perform the steps of preprocessing the sequence on the basis of a so-called content-change strength (CCS) computed for each frame, generating a predicted frame using said motion vectors estimated for each block, applying to a difference signal between the current frame and the last predicted frame a transformation sub-step producing a plurality of coefficients and followed by a quantization sub-step of said coefficients, and coding said quantized coefficients. According to the invention, said CCS is used in said quantization sub-step for modifying the quantization factor, said CCS and the quantization factor increasing or decreasing simultaneously.

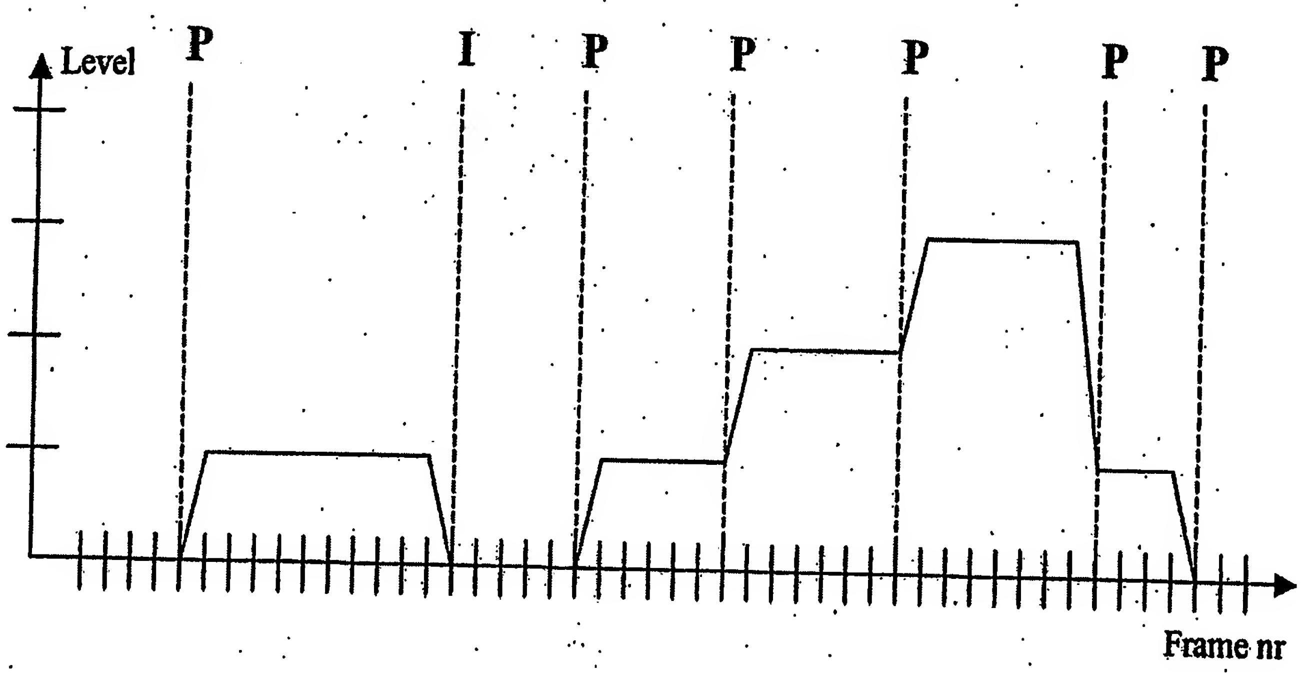
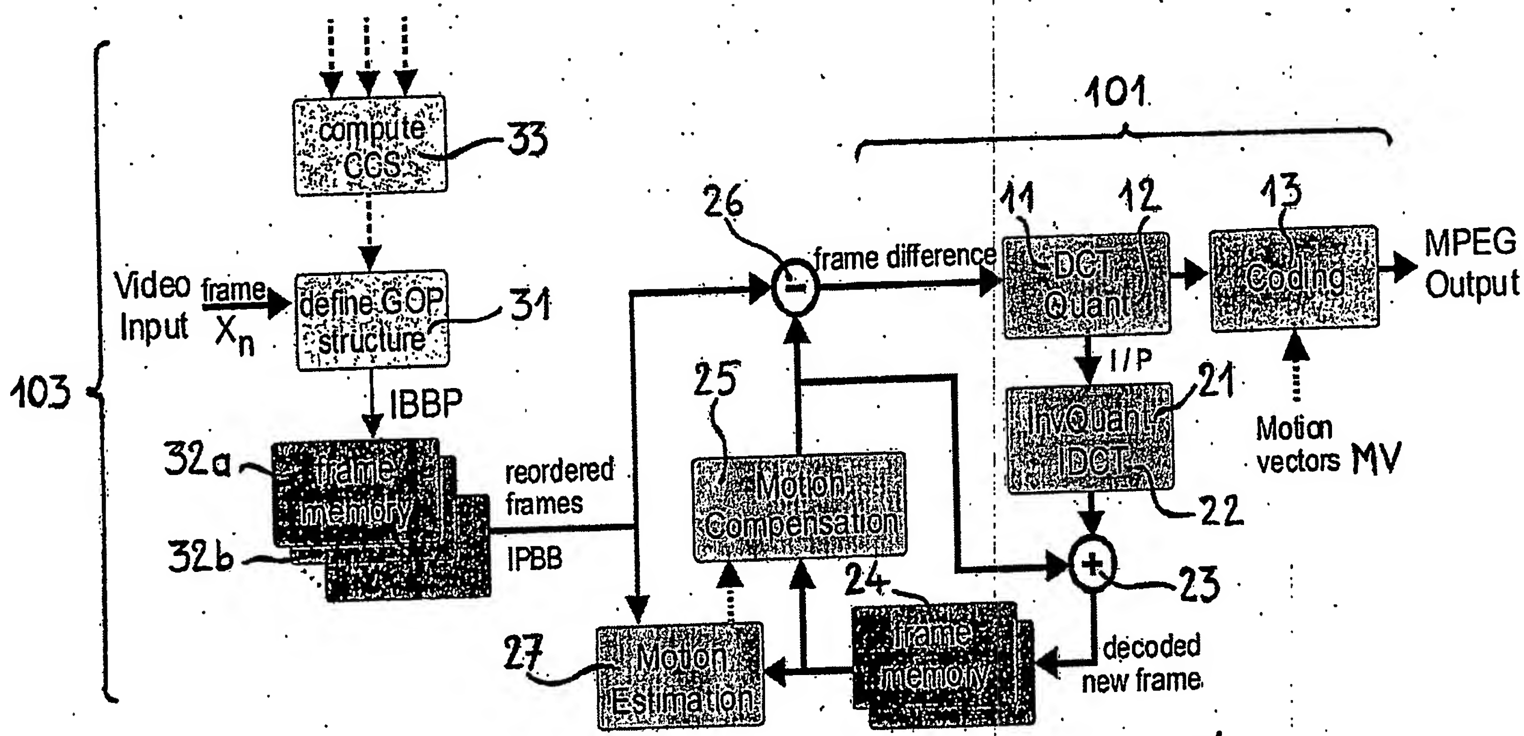


FIG.1



compute CGS
33
define GOP structure
31
Video Input
frame
X_n
IBBP
103
32a
frame memory
32b
reordered frames
IPBB
26
frame difference
25
Motion Compensation
27
Motion Estimation
24
frame memory
102
101
11
12
DCT Quant
13
Coding
MPEG Output
I/P
21
InvQuant IDCT
22
23
decoded new frame
Motion vectors MV

Fig.2

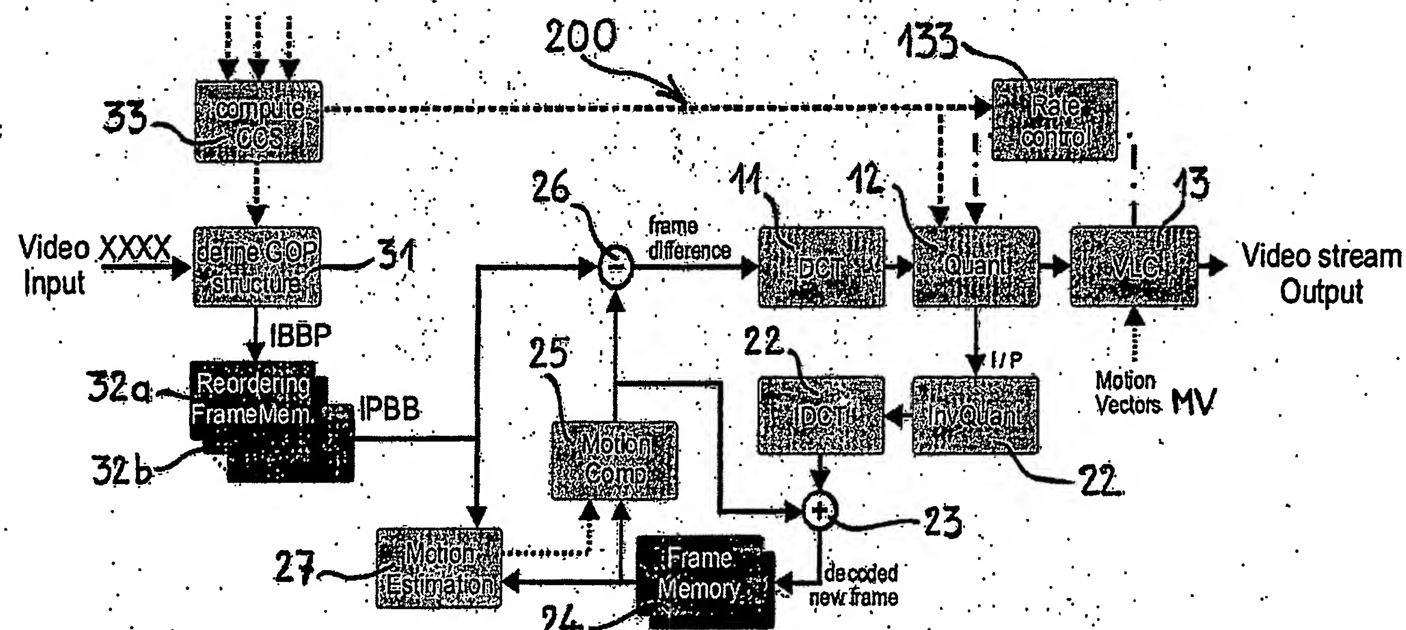

compute CCS
33
200
133
Rate control
define GOP structure
31
Video XXXX Input
IBBP
IPBB
Reordering FrameMem.
32a
32b
26
frame difference
11
DCT
12
Quant
13
VLC
Video stream Output
I/P
Motion Vectors MV
22
IDCT
InvQuant
25
Motion Comp
23
decoded new frame
27
Motion Estimation
Frame Memory
24

FIG.3

PCT/IB2004/003618

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.